AD-A013 572

NONLINEAR FILTERING

R. S. Bucy

University of Southern California

Prepared for:

Air Force Office of Scientific Research

July 1975

DISTRIBUTED BY:

NTIS

National Technical Information Service
U. S. DEPARTMENT OF COMMERCE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ADA013572

234120

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR - TR - 75 - 1163	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) NONLINEAR FILTERING		5. TYPE OF REPORT & PERIOD COVERED Interim
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) R. S. Bucy		8. CONTRACT OR GRANT NUMBER(s) AFOSR-71-2144
9. PERFORMING ORGANIZATION NAME AND ADDRESS University of California, Department of Aerospace Engineering, Los Angeles, CA 90007		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 61102F 9769-01
11. CONTROLLING OFFICE NAME AND ADDRESS Air Force Office of Scientific Research/NM, 1400 Wilson Blvd, Arlington, VA 22209		12. REPORT DATE July 1975
		13. NUMBER OF PAGES 6
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear Filtering
Monte Carlo
Phase Demodulation
Representation Theorem
Convolution

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

A review of the work of numerical realization of Optimal Nonlinear Filters is given as well as a historical account of the development of the theory. The problems of synthesis and numerical representation of the condition density of the signal given the observation are treated in detail.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

NONLINEAR FILTERING[1]

R. S. Bucy

Departments of Aerospace Engineering and Mathematics
University of Southern California
Los Angeles, Calif. 90007

Abstract

The purpose of this paper is to review recent progress in the actual construction of optimal nonlinear filters. Specifically, we will consider the construction of numerical algorithms, which accept as inputs noisy observations of a nonlinear function of signal process and produce as outputs estimates of the signal process. These algorithms, because of the structure of the nonlinear filtering problem, can be thought of conditional probability function generators. A historical review of the theory and applications of nonlinear filtering will be given, which will attempt to catalog the seminal ideas in the field as well as some unsolved problems which are obstructing progress.

1. Introduction

History of the Theoretical Resolution

The theory of nonlinear filtering was initiated by Stratonovich in [1] in 1960. He assumed that the signal $x(s,w)$ was a Markov diffusion process taking values in R^d, and further that observations were available of the form

$$z(t,w) = \int_{t_0}^{t} h(s,x(s))\, ds + v(t,w) \tag{1.0}$$

where $v(t,w)$ is a Brownian motion process R^s valued with infinitesimal covariance $R(t)$ and h is measurable function from $(t_0, \infty) \times R^n$ to R^s. The problem consisted of finding the conditional distribution of $x(t)$ given $z(s,w)$ for $t_0 \le s \le t$, as if one is interested in estimating $x(t)$ on the basis of knowledge of the observation sample function $z(\cdot,w)$, this conditional distribution contains all the information relevant to the estimation process. Let us assume that $x(t)$ possesses a transition density

$$p(t,x,y) \text{ (i.e., } P(x(t) \in A|x(o) = x) = \int_A p(t,x,y)\, dy \text{)}$$

and further the following Lindeberg conditions hold

$$E(x(t+h) - x(t))|x(t) = y) = f(y)h + o(h) \tag{1.1}$$

$$E\Big((x(t+h) - x(t))(x(t+h) - x(t))'|x(t) = y\Big) = \sigma(y)\, Q\, \sigma'(y)\, h + o(h)$$

Under these conditions and smoothness Kolmogorov in [2] demonstrated that

$$\frac{\partial p}{\partial t} = Ap \tag{1.2}$$

where $A = f \cdot \frac{\partial}{\partial x} + \frac{1}{2} \operatorname{trace} \sigma Q \sigma' \left[\frac{\partial^2}{\partial x_i\, \partial x_j}\right]$,[2] A the so-called backward equation. The equation defines the evolution of the expectation of functions, the conditional density is a solution of (1.1) with initial value $p(o,x,y) = \delta(x-y)$, or in other words $p(t,x,y)$ is the Green's function of (1.1). If on the other hand the initial value is $k(y)$, the solution is $E\, k(x(t))|x(o) = x)$. A dual equation was also derived as

$$\frac{\partial \lambda}{\partial t} = \tilde{A} \lambda \tag{1.3}$$

where $\tilde{A}$ is the formal adjoint of A and the general solution of it is

$$\lambda(t,x) = \int p(t,z,x)\mu(z)dz$$

or probabilistically $\lambda(t,x)$ is the density of $x(t)$ when $x(o)$ has density μ. Notice that $p(\cdot,\cdot,y)$ satisfies (1.1) while $p(\cdot,x,\cdot)$ satisfies (1.2). Suppose $k(x,t,z_\cdot)$ is the conditional density of $x(t)$ given $z(s,w)$ $t_0 \le s \le t$, then Stratonovich showed, using Bayes rule and truncating a Taylor series, the equivalent of[3]

$$dk = \tilde{A}\, k dt + (h-\hat{h}_t)' R^{-1}\, d\, I\, k \tag{1.4}$$

where (1.3) is a random differential equation of Ito type, see [3]. The initial condition of (1.4) is the density μ of $x(o)$, and $\hat{h}_t = \int h(t,y)\, k(t,y)\, dy$ with $dI = dz - \hat{h}\, dt$. In [6], Ito showed that if (1.1) holds, f and σ are Lipshitz and then $x(t)$ itself is the solution of the Ito equation

$$dx = f(x)dt + \sigma(x)d\beta \tag{1.5}$$

where β is a Brownian vector process independent of $x(o)$ with infinitesimal spectral matrix Q.

Now (1.4) is locally a description of the nonlinear filtering problem, and since in Quantum Mechanics the local description, the Schroedinger equation, has a global analog the famous Feynman Path Integral, see [38] and [39], one might ask, does the nonlinear filtering problem possess a global description? In fact, this is the case, although it was not until 1965 in [7] that this global description, The Representation Theorem, was proposed. Assuming regularity conditions, see [8] through [14], for the details[4]

$$k(t,x) = \frac{E^{z\cdot}\, e^{H_t}|x(t) = x)\, \mu(x)}{E^{z\cdot}\, e^{H_t}} \tag{1.6}$$

where $E^{z\cdot}$ means average with the observation path fixed and

$$H_t = \int_{t_0}^{t} h(s,x_s) R_s^{-1}\, dz - \frac{1}{2} \int_{t_0}^{t} \|h(s,x_s)\|^2_{R_s^{-1}}\, ds \tag{1.7}$$

It is interesting to note that the conditions for the

[1] This research was supported in part by the United States Air Force, Office of Scientific Research, Air Force Systems Command, under AFOSR Grant 71-2144.

[2] The order within the trace is important here.

[3] Actually, in [1] another form of (1.3) is given which requires to be valid in interpretation in terms of a new integral detailed in [4]. Kushner in [5] was the first to derive (1.3) in the Ito form. See also [33].

[4] Mortensen in [9] first recognized that the representation theorem was in fact derivable from the chain rule for Radon-Nikodym derivatives in function space; the most elegant proof so far is given in [12], where a Hilbert space setting reduces the nonlinear problem to a linear one, where the result is easy.

validity of (1.6) are conditions on the signal process and the sensor or conditions, not á priori unverifiable conditions on k such as; k be twice continuously differentiable which is necessary for (1.4) to be valid.

A largely heuristic approach to nonlinear filtering theory, the so-called innovations approach, discovered by Frost in [15] and popularized by Kailath and Frost in a number of papers (see [16] for references), hinges on transforming the observations to produce a new observation process that is white and consists of "new information" at each instant, generalizing ideas of Kolmorogorov in [17]. While these ideas are clearly useful, a number of results[5] claimed have yet to be proven.

It seems convenient to consider here the discrete sequential version of the representation theorem as for numerical purposes it seems the most useful--see [17] for an early occurrence of this result. Suppose both $z(t,w)$ and $x(t,w)$ are sampled with sampling interval Δ and denote

$$x_n = x(n\,\Delta + t_o,w)$$

$$z_n = z(n\,\Delta + t_o,w) - z((n-1)\,\Delta + t_o,w)$$

Further, suppose that the conditional density of $x_{n+1} = y$ given $x_n = x$ is $S_n(y,x)$ and the conditional density of z_n given $x_n = x$ is $D_n(x,z_n)$ then

$$P_{n+1}(y) = \int_X S_n(y,x)\, F_n(x)\, dx^{6} \tag{1.8}$$

$$F_n(x) = \gamma_n\, D_n(x,z_n)\, P_n(x) \tag{1.9}$$

where $P_n\{F_n\}$ are respectively the conditional densities of x_n given $z_{n-1} \ldots z_1$, $\{z_n, z_{n-1} \ldots z_1\}$, and γ_n is the appropriate normalizing term. Note that (1.8) represents model following while (1.9) represents the influence of the new piece of data, the analog of the contributing factors of estimate dynamics in the linear case.

2. Problems Arising in Numerical Realization

Let us note that in continuous time both the local and global dynamics of the conditional density (1.4) and (1.6) involve a non-pointwise limiting process, specifically a limit in the mean because of the definition of the stochastic integral--see [6]. In view of this finding, the value of $k(t,x,z_\cdot)$ when a sample path $z(\cdot,w)$ is given by direct difference approximation of (1.4) or replacing $x(t,w)$ in (1.6) by random process which has at most finite number of values for each w can lead to divergent approximations negative values for the approximations to the density $k(t,x)$ in the case of (1.4) and in general disasterous numerical behavior.

It is also a problem, illustrative of our last remarks, to produce numerically the continuous time white noise processes sample functions, in fact, it was shown by Wong and Zakai in [18] that the solution of the scalar stochastic differential equation

$$dx = f(x)\, dt + \sigma(x)\, d\beta$$

is in general different from the limit of

$$\frac{dx_n}{dt} = f(x_n) + \sigma(x_n)\, W_n \tag{2.1}$$

where W_n is the derivative of an absolutely continuous function and such that

$$\beta(t,w) - \beta(\tau,w) = \lim \int_t^t W_n(s)\, ds$$

in fact, $x^* = \lim x_n$ satisfies

$$dx^* = f(x^*)dt + \frac{d\sigma}{dx^*}\,(x^*)dt + \sigma(x^*)d\beta \tag{2.2}$$

On the other hand, numerous procedures for computer realization of approximate white noise sequences exist, although most of them are fairly poor approximations, especially the canned subroutines available for the IBM and CDC machines, and most of the others pass statistical tests which depend fundamentally on their assumed ergodicity. In [19], Senne develops a generator which is not only machine-independent[7] but further passes the Kolmogorov-Smironov test for distributional fit. For all generators judicious choice of the seed is important.

It appears then that it is preferable to sample both the signal, $x(t,w)$ and $z(t,w)$ at a rate faster enough not to lose information relative to the continuous problem--see [20] and [21] for an analysis which determines the sampling rate for the phase demodulation problem-- and to use (1.8) and (1.9) to realize the nonlinear filter.

Another numerical problem is that in order to evaluate filter performance, Monte Carlo runs must be performed. This requirement taxes the ability of modern third generation digital computers for problems with low state dimensions signal processes. Further, the number of Monte Carlo repetitions must be large enough to provide small enough confidence bands on the error performance so that the optimal filter performance can be meaningfully compared with sub-optimal[8] filters--see [22]. Hopefully, more research on á priori bounds will eliminate the need for costly Monte Carlo simulations. Promising research in this direction is reported in [23].

Finally, I think it is appropriate at this point to indicate why it is important to undertake numerical realization studies. Primarily these studies are important in order to conclusively demonstrate the degree of superior error performance which can be achieved using the optimal nonlinear filter. A subsidiary benefit is that insight is gained on the behavior of nonlinear filters. Because of a paucity of examples where closed form solutions exist--see [24] for a number of such examples--there are few opportunities to check conjectures as well as to gain insight into what properties might be generally true. Without examples, the field of nonlinear filtering could easily develop into a stale effete area which dies by feeding on itself and is overburdened by work which is neither good mathematics nor useful engineering.

3. Conditional Density Representation

It is clear that for digital computer iteration of (1.8) and (1.9) a map T from a subset L of $[o,1]^X$ to a finite dimensional vector space K must be given--here χ is a subset of R^d. If χ is compact, the map can be fixed, while if χ is not compact, the map must change with time. Some examples will clarify the general idea.

Example 1 (see [25])

$$L = C_0(-\infty,\infty)\ ,\quad d = 1$$

[5] Specifically the proof of the innovations theorem in [15] is wrong.
[6] All variables can take vector values and the integral may be multi-dimensional.
[7] Dr. Senne informs me that the generator has been realized on an H.P.65 programable hand-calculator, so that base 2 assumptions in [19] are unnecessary.
[8] Notice that generally error performance of sub-optimal filters must be evaluated by Monte Carlo methods. Further, the statistical design of the Monte Carlo trials mustallow for the nonergodic nature of the error.

$$(T_n P_n)(x) = \left\{\tilde{P}_n(x_i^n)\right\}_{i=1\,..\,2M+1} \quad \tilde{P}_n(x)$$

is the average of P_n over a ball-centered at x.

$$x_i^n = \mu_{n|n-2} + \frac{\sigma_{n|n-2}}{2M+1}\,(i - M - 1)$$

where M is an integer and $\mu_{n|n-2}$ and $\sigma^2_{n|n-2}$ are

$$\mu_{n|n-2} = \tilde{E}\, x_n | z_0 \,\ldots, z_{n-1})$$

$$\sigma^2_{n|n-2} = \tilde{E}(x_n^2 | z_0, \ldots, z_{n-1}) - \mu^2_{n|n-2}$$

and the $\tilde{}$ superscript denotes averaging with respect to

$$T^{\#}_{n-1}\,(T_{n-1} P_{n-1})(x) = \sum_{j=1}^{2M+1} \tilde{P}_{n-1}(x_j^{n-1})\,\delta(x - x_j^{n-1}).$$

$T^{\#}_{n-1}$ denote a choice of pre-image of an element in the range of T_{n-1}.

Example 2 $d = 2$ (see [29])

$$(T_n P_n) = \left\{\tilde{P}_n(x_i^n, y_j^n)\right\}_{i=1\ldots 2M+1,\ j=1\ldots 2N+1}$$

$\tilde{P}_n(x,y)$ denotes the average of P_n over a ball of small radius centered at (x,y)

$$\begin{pmatrix} x_i^n \\ y_j^n \end{pmatrix} = \mu_{n|n-2} + \frac{\lambda_1^n}{2M+1}(i-M-1)\underline{e}_1^n + \frac{\lambda_2^n}{2N+1}(j-N-1)\underline{e}_2^n$$

M and N are integers

λ_i^n, $\underline{e}_i^n$ are eigenvalue and eigenvectors of $S_{n|n-2}$

where

$$S_{n|n-2} = \tilde{E}\, x_n\, x_n' \,|z_0, \ldots, z_{n-2}) - \mu_{n|n-2}\,\mu'_{n|n-2}$$

$$\mu_{n|n-2} = \tilde{E}\, x_n | z_0 \,\ldots, z_{n-2})$$

and $\tilde{}$ superscript denote averaging with respect to the density

$$\sum_{i=1}^{2M+1} \sum_{j=1}^{2N+1} \tilde{P}_{n-1}\,(x_i^{n-1}, y_j^{n-1})\,\delta(x - x_i^{n-1})\,\delta(x - y_j^{n-1}).$$

Example 3 L is the set of continuous probability densities on the Torus

$$T\,P_n = \left\{a^n_{\ell_1, \ell_2, \ldots \ell_r}\right\}_{|\ell_j| \le N_j}$$

N_j are integers, and $a^n_{\ell_1 \ldots \ell_r}$ are Fourier coefficients of P_n (see [27]).

Example 4 (see [28])

The map T assigns to a function a finite subset, its interpolative spline under tension coefficients.

Example 5 (see [49])

The map T assigns to a function a finite subset of its coefficients in a Gauss-Hermite expansion.

Example 6 (see [20], [26])

L is the set of functions on the Torus in R^d and T assigns to a function its values on a uniform grid of meshes

$$\delta_{\ell_1},\ \delta_{\ell_2},\ \delta_{\ell_d}$$

in each coordinate.

Example 7 (see [30])

The map T assigns to a function a finite subset of its non-interpolative spline coefficients.

Example 8 (see [31])

The map T assigns to a function coefficients of a least squares or L^2 fit of the function to a finite linear combination of functions.

In all of these cases (1.8) and (1.9) are approximated for synthesis purposes by the vector matrix recursion relation

$$J_n \doteq K(n)\, J_{n-1} \tag{3.1}$$

where J_n is the image of either $F_n(x)$ or $P_n(x)$ [9] under T_n and $\doteq$ indicates that J_n must be renormalized or transformed so that a canonical choice of pre-image of J_n, which we denote by $T^{\#} J_n$, is a density. The relation (3.1) can be arrived at in the following way. First, one notes that in the one-step predictor case, for example

$$P_{n+1}(x) \doteq \gamma_n \int_X S(x,y)\, D_n(y, z_n)\, P_n(y)\, dy$$

$$\doteq \gamma_n \int_X S(x,y)\, D_n(y, z_n)\, T^{\#}_n\, T_n[P_n(y)]\, dy$$

and applying T_{n+1} to both sides, it follows that

$$J_{n+1} = T_{n+1}\left\{\gamma_n \int_X S(x,y)\, D_n(y, z_n)\, T^{\#}_n J_n(y)\, dy\right.$$

which is equivalent to (3.1). A problem which leads to numerical instability is the following; suppose (T_n) and $(T^{\#}_n)$ are chosen and the relation

$$J_n = K(n)\, J_{n-1} \tag{3.2}$$

is iterated, the sequence $T^{\#}_n J_n$ does not always remain positive, even when J_0 is a vector with $T^{\#}_0 J_0$ positive. In Example 5, a convenient and effective modification of (3.2) to preserve positivity is redefining J_n as

$$J_n(i) = \max\,(o, (K(n) J_{n-1})(i))$$

In examples 1, 2 and 6, $T^{\#}_n$ can be chosen so that the above negativity effect does not arise.

For problems where the signal process is a degenerate random process (i.e., a random variable), the representation theorem gives an explicit expression for the conditional density and the problem of representing the density is trivial--see [32] for results concerning this degenerate case.

The point mass representation, Examples 1 and 2, was the first one considered and, in fact, can be made

[9] The choice depends on whether one wishes to synthesize the filter or one-step predictor.

quite accuracy by increasing the number of grid points until the signal estimates for a fixed sequence of observations agree to say four places by successive choices of finer subdivisions of the grid. The accuracy obtained by this method is not quite unexpected--see for example the discussion on Page 4 of [34], where coincidence of form is compared with metric closeness. The drawback of the point mass method consists of the large computation time per estimate, and in fact the other representation methods were motivated by the desire to decrease this estimate time, while preserving a given accuracy relative to a point mass accuracy benchmark. A somewhat different approach to the representation problem consist of determination of a perturbation series for P_n and F_n in (1.8) and (1.9) when $S_n(y,x)$ depends on a parameter q; for example, suppose

$$S_n(x,y) = \frac{1}{\sqrt{2\pi q}} e^{-\frac{(x-y)^2}{2q}}.$$

then F_n and P_n can be determined as series in q, see [35] and [36] for complete details. This latter approach is numerically investigated in [36].

The representation problem is quite important in that the time between estimates can be improved by an order of magnitude through a careful choise of the representation. While clearly this problem of representation is important and deserves careful study, it seems to be a second order effect relative to computation time of estimates, while the choice of synthesis device is first order. In a later section we will discuss other synthesis devices which promise two or more orders of magnitude speed improvement over synthesis by third generation serial digital computers.

4. A Typical Problem

The problem of phase demodulation is a problem of low state dimension and has been extensively investigated both from the point of view of optimal and suboptimal design--see [20] for references to a universally used suboptimal design, the phase lock loop. For this problem the following model is appropriate:

$$\begin{aligned} x^1_{n+1} &= x^1_n + \Delta x^2_n \\ x^2_{n+1} &= x^2_n + u_n \end{aligned} \tag{4.1}$$

where u_n is a Gaussian white noise sequence of zero mean and variance Δq. The initial condition on (4.1) is bivariate normal and independent of the plant noise u_n. The observation process is

$$\begin{aligned} z^1_n &= \cos x^1_n + v^1_n \\ z^2_n &= \sin x^1_n + v^2_n \end{aligned} \tag{4.2}$$

where v^i_n are independent Gaussian white noise sequences of zero mean and variance r/Δ, and uncorrelated with x^1_0, x^2_0 and u_n. The sampling rate Δ is chosen as

$$\Delta = .1\sqrt{2}\left(\frac{r}{q}\right)^{\frac{1}{4}},$$

on the basis of a linear analysis to assure good approximation of continuous data--see [21] and [26].

If the sensor were linear, the Wiener theory would show that the mean square error in estimating phase x^1_n is

$$R = \sqrt{2}\, q^{\frac{1}{4}} r^{3/4}$$

for continuous observations, which, of course, is a lower bound on the mean square error of the phase demodulation problem.

The cyclic loss function, $1/2(1 - \cos(x^1_n - x^*_n))$ was considered originally in [4] and rediscovered in [40] and still later in [41], all in the context of a less realistic phase demodulation problem where the phase is Brownian motion[10], instead of the integrated Brownian motion model represented by (4.1). This cyclic loss function is appropriate for problems where one is interested only in estimating relative phase. The cyclic estimate that x^*_n, which minimizes the cyclic loss, is the argument of

$$a^n_{-1,0} = \frac{\Delta}{(2\pi)^2}\int_0^{2\pi}\int_0^{2\pi/\Delta} e^{ix}\, \mathcal{J}_n(x,y)\,dx\,dy ,$$

where $\mathcal{J}_n(x,y)$ is the conditional distribution of x^1_n mod 2π, x^2_n mod $2\pi/\Delta$ given the observations. In fact, by consideration of the estimation of relative phase, the relevant conditional distribution for filtering can be taken as the distribution on Torus, T arising from conditional distribution of phase and phase rate, x_n, x^2_n given the observations, $\mathcal{J}_n(x,y)$, where

$$\mathcal{J}_n(x,y) = \sum_{\mu,\lambda} \mathcal{J}_n\left(x + 2\pi\nu,\ y + \frac{2\pi}{\Delta}\lambda\right) \tag{4.3}$$

for $(x,y) \in T$. Extensive Monte Carlo simulation of the cyclic nonlinear filter has shown that the cyclic estimate achieves a 3-db error performance improvement over the phase lock loop--see [42]. The first results were obtained for a single q and a point mass filter, see example 6 of Section 3, and for each value of R, three hours of 6600 C.P.U. were required. Later, by using the Fourier representation of the density, the C.P.U. time was cut by a factor of 10. Finally, in [43] we demonstrated that mean square error for the optimal demodulator was independent of q. Details on the representation, Example 7, Section 3. can be found in [44].

It is clear from this example that, while significant error variance reduction is possible with a serial digital computer as a realization device, the massive computational task associated with accurate synthesis and Monte Carlo error analysis limit the state dimension of the nonlinear filters one can effectively build and analyze.

5. More Effective Synthesis Devices

It became clear very early that serial realization was effectively speed-limited by the convolution task, (1.8), necessary to "update" the a priori conditional density to obtain the a posteriori conditional density for the state Δ seconds later, when a new piece of data is received. From the structure of (1.8), it is clear that immense estimate computation time reduction can be obtained by using a parallel digital computer as the synthesis tool. An analysis of the possible savings is given in [29].

[10] When the phase is Brownian motion, error variance improvement due to using the nonlinear filter is only about .7 db, and further, the absence of the necessity of phase rate tracking makes the problem of little practical interest, except perhaps for classroom discussion.

A feasibility study of the synthesis of optimal filters using a hybrid system to achieve parallelism for the convolution task is reported in [46]. This study used a serial machine to simulate a contemporary hybrid system with MOBSSL as the simulation language. The results obtained in this feasibility study indicated that a hybrid system was capable of achieving considerable time saving, albeit with only two place estimate accuracy. In the last year, a hybrid nonlinear filter was built at the Labatorio d'Automatico, University Polytecnico Barcelona, Spain, using an Electronic Associates EAI-680 hybrid system with a floating point processor. This hybrid nonlinear filter achieved the characteristics predicted in [46], and the results are reported in [45].

Another approach is using a contemporary parallel machine, say the Illiac, as the synthesis tool; preliminary estimates indicate that three-hour Monte Carlo runs on the CDC 6600 can be accomplished in three minutes on the Illiac and, more importantly, nonlinear filters corresponding to problems with four-state dimensional signal process models can be built and Monte Carlo error analysis performed routinely. This is an area of our current research interest.

Finally, it is clear that special purpose serial machines fabricated on acoustic-optic or surface wave principles in theory and for simple signals in practice can achieve temporal convolutions in 6600 cycle time about 200 nanoseconds--see [47]. In [48], an approximate homomorphism between the Banach Algebra B_1 of periodic functions of one variable and the Banach Algebra B_r of functions on the r-dimensional torus. The multiplication in these algebras is the appropriate convolution. When $r = 2$, then F and G in B_2, we have

$$\phi(F*G) \doteq \phi(F)*\phi(G) \tag{5.1}$$

$$\phi : B_2 \to B_1$$

where ϕ is the appropriate ring homomorphism. The meaning of (5.1) is that (1.8) can be computed by performing a temporal convolution of

$$\phi(S_n) \quad \text{and} \quad \phi(F_n) \quad ,$$

for the phase demodulation problem and the temporal convolution can be done using surface waves generated by $\phi(F_n)$ on a piezo-electric crystal with photographically deposited metallic fingers corresponding to $\phi(S_n)$. This becomes most interesting when S_n is independent of n as in the case of the phase demodulation problem. Such a device is currently in the planning stage and is a joint research project of the author and Dr. Eugene Dieulesaint of Ecole Superieur de Chieme et Physie, Paris.

6. Conclusions

In this paper we have reviewed some of the first attempts to synthesize the optimal nonlinear filter. The problem itself, while extremely important, induces solution methods which are extremely time consuming because of what Bellman has aptly called the curse of dimensionality. The current technics, while primative, are applicable to a wide variety of problems, including, for example, the solution of parabolic partial differential equations in more than one space dimension and are of importance if only for this application. This survey will have served its purpose if it succeeds in interesting research workers in pursuing these problems further and developing new methods of practical synthesis.

7. References

[1] R. L. Stratonovich, "Conditional Markov Processes," Theory of Prob. Appl., 5, 1960, 156-178.

[2] A. N. Kolmogorov, "Über die Analytischen Methoden in der Wahrscheinlichkeitsrechnung," Math. Ann., 104, 1931, 415-458.

[3] K. Ito, "On Stochastic Differential Equations," Mem. Amer. Math. Soc., Vol. 4, 1951.

[4] R. L. Stratonovich Conditional Markov Processes and Their Application to Optimal Control Elsevier, New York, 1966.

[5] H. P. Kushner, "On the dynamical equations of conditional probability density functions, with applications to optimal stochastic control theory," J. Math. Anal. Appl., 8, 1964, 332-344.

[6] K. Ito, "Lectures on stochastic processes," Lecture Notes, Tata Institute of Fundamental Research, Bombay, India, 1961.

[7] R. S. Bucy, "Nonlinear filtering theory," I.E.E.E. Trans. Automatic Contr. (corresp.), AC-10, 1965 198.

[8] R. S. Bucy and P. D. Joseph Filtering for stochastic processes with applications to guidance Interscience, New York, 1968.

[9] R. E. Mortensen, "Optimal control of continuous time stochastic systems," Thesis, E.E. Dept., University of California, Berkeley, 1966.

[10] R. S. Bucy, "Linear and nonlinear filtering," Proc. I.E.E.E., 58, 1970, 854-864.

[11] G. Kallianpur and C. Striebel, "Estimation of stochastic systems, etc." Annals of Math. Stat., 39, 1969, 785-801.

[12] A. Bensoussan, "Statistical estimation in Hilbert spaces: Application to Bucy's representation theorem," Proc. 6th Prague Sym. on the Formation Theory, 1971, Academia, Prague, 1973, 107-123.

[13] P. A. Meyer, "Sur un probléme de filtration," Seminar de probabilities, Pt VII, Lecture Notes in Math., 321, Springer, New York, 197-206.

[14] YuSkorokhod Studies in the theory of random process Addison Wesley, Reading Mass, 1965.

[15] P. A. Frost, "Estimation in continuous-time nonlinear systems," Thesis, Dept. of Electrical Engineering, Stanford University, 1968.

[16] T. Kailath, "A view of three decades of linear filter theory," I.E.E.E. Trans. on Inform. Theory, IT-20, 2, 1974, 146-180.

[17] Y. C. Ho, R. C. K. Lee, "A Bayesian approach to problems in stochastic estimation and control," I.E.E.E. Trans. Auto.Contr., 9, 1964, 333-339.

[18] E. Wong, M. Zakai, "On the relations between ordinary and stochastic differential equations and applications to stochastic problems in control theory," Proc. 3rd IFAC Congr.,London: Butterworth, 1966.

[19] K. D. Senne, "Machine independent Monte Carlo evaluation of the performance of dynamic stochastic systems," Stochastics, 1, 3, 1974, 215-238.

[20] R. S. Bucy, A. J. Mallinckrodt, "An optimal phase demodulator," Stochastics, 1, 1, 1973, 3-23.

[21] R. S. Bucy, C. Hecht, K. D. Senne, "New methods for nonlinear filtering," Rev. Francais d' Automatique, J-1, 1973, 3-54.

[22] R. S. Bucy, "Realization of nonlinear filters," Proc. 2nd Symp. on Nonlinear Estimation, 1971, Western Periodicals.

[23] M. Zakai and J. Ziv, "Lower and upper bounds on the optimal filtering error of certain diffusion processes," I.E.E.E. Trans. on Inform. Theory, IT-18, 1972, 325-331.

References (continued)

[24] J. T. Lo, "Finite dimensional sensor orbits and optimal nonlinear filtering," I.E.E.E. Trans. Inform. Theory, IT-18, 5, 1972, 583-588.

[25] R. S. Bucy, "Bayes theorem and digital realization for nonlinear filters," XVII, 2, 1969, 80-94.

[26] R. S. Bucy, C. Hecht, K. D. Senne, "An engineer's guide to building nonlinear filters," F. J. Seiler Res. Lab. Rt. SRL-TR-72-0004, U.S. Air Force Academy, Colorado Spring, 1972.

[27] R. S. Bucy, H. Youssef, "Fourier realization of the optimal phase demodulator," Proc. 4th Symp. on Nonlinear Estimation Theory and Its Applications, San Diego, Western Periodicals, 1973, 34-38.

[28] H. Youssef, "Interpolative spline filters," PhD.Thesis, Aerospace Eng. Dept., University of Southern California, 1975.

[29] R. S. Bucy, K. D. Senne, "Digital synthesis of nonlinear filters," Automatica, 7, 1971, 287-298.

[30] R. J. P. de Figueiredo and Y. G. Jan, "Spline filters," Proc. 2nd Symp. Nonlinear Est. Theory and Applications, San Diego, 1971, Western Periodical, 88-99.

[31] H. W. Sorenson and D. L. Aslpach, "Recursive Bayes estimation using Gaussian sums," Automatica, 7, 1971, 465-479.

[32] B. W. Licht, Thesis, Systems Research Center, Case Institute of Tech., Cleveland, 1970.

[33] G. Kallianpur, C. Striebel, "Stochastic differential equations occurring in the estimation of continuous parameter stochastic processes," Theory Prob. and Appl., 16, 5, 1969, 597-622.

[34] Rene Thom, "Structural stability and morphogenesis," Benjamin, Reading, Mass., 1975.

[35] GeraldAlengrin, Thesis d'Etat, University Paul Sabatier, Toulouse, France, 1974.

[36] R. Tye, Thesis, Dept. of Aerospace Eng., University of Southern California, Los Angeles 1975.

[37] R. S. Bucy, R. Tye, H. Youssef, "Nonlinear filtering and singular perturbation," Proc. 6th Symp. on nonlinear estimation and its Applic., San Diego, 1975.

[38] J. B. Keller and D. W. McLaughlin, "The Feynman integral," Amer. Math. Monthly, 82, 5, 1975, 451-465.

[39] R. P. Feynman, A. R.Hibbs, Quantum Mechanics and Path Integrals, McGraw-Hill, New York, 1965.

[40] A. J. Mallinckrodt, R. S. Bucy, S. Y. Cheng, "Final Report for a design study for an optimal nonlinear receiver/demodulator," NASA Contract: NAS5-10789 Goddard Space Flight Center, Greenbelt, Maryland, August 31, 1970.

[41] A. S. Willsky, "Fourier series and estimation,etc." Proc. Inf. Theory, IT-20, 1974, 577-583.

[42] R. S. Bucy, H. Youssef, "Optimal phase demodulation," submitted: I.E.E.E. Trans. Autom. Contr.

[43] R. S. Bucy, H. Youssef, "Dependence of the optimal phase demodulator on statistical parameters," I.E.E.E. Trans. Autom. Contr., AC-20,2, 1975, 259-260.

[44] H. Youssef, "Interpolative spline filters," Thesis, Dept. of Aerospace Eng., University of Southern California, Los Angeles, Jan. 1975.

[45] L. Basanez, P. Brunet, R. S. Bucy, R. Huber, D. S. Miller and J. Pages, "Hybrid nonlinear filters," submitted: Proc. 6th Symp. Nonlinear Estimation and Appl., San Diego, 1975.

[46] R. S. Bucy, M. Merrit, D. S. Miller, "Hybrid synthesis of the optimal discrete nonlinear filter," Stochastics, 1, 1973, 151-211.

[47] E. Dieulesaint, D. Royer, Ondes Elastique dans les Solides, Avec applications traitement du signal, Masson, Paris, 1974.

[48] R. S. Bucy, A. J. Mallinckrodt, H. Youssef, "High speed convolution of periodic function," submitted: Siam Jour. Appl. Math.

[49] C. Hecht, "Digital realization of nonlinear filters," Proc. 2nd Symp. on Nonlinear Estimation Theory and Its Appl., San Diego, 1971, 88-99; see also: Thesis, Aerospace Eng. Dept., University of Southern California, Los Angeles, 1972.